



Age-, sex- and ethnicity-related differences in body weight, blood pressure, HbA_{1c} and lipid levels at the diagnosis of type 2 diabetes relative to people without diabetes

Alison K. Wright^{1,2} · Paul Welsh³ · Jason M. R. Gill³ · Evangelos Kontopantelis⁴ · Richard Emsley⁵ ·
Iain Buchan⁶ · Darren M. Ashcroft² · Martin K. Rutter^{1,7} · Naveed Sattar³

Received: 19 February 2020 / Accepted: 2 April 2020 / Published online: 21 May 2020
© The Author(s) 2020

Abstract

Aims/hypothesis The aim of this work was to determine how weight patterns together with blood glucose, BP and lipids vary at diagnosis of diabetes by age, sex and ethnicity.

Methods Using the UK Clinical Practice Research Datalink, we identified people with type 2 diabetes ($n = 187,601$) diagnosed in 1998–2015 and compared their weights, HbA_{1c}, BP and lipid levels at diagnosis with age-matched people without diabetes ($n = 906,182$), by sex and ethnic group.

Results Younger age at diagnosis was associated with greater adjusted mean difference (95% CI) in weight between those with vs without type 2 diabetes: 18.7 (18.3, 19.1) kg at age 20–39 years and 5.3 (5.0, 5.5) kg at age ≥ 80 years. Weight differentials were maximal in white women, and were around double in white people compared with South Asian and black people. Despite lower absolute values, BP differences were also greater at younger age of diabetes onset: 7 (6, 7) mmHg at age 20–39 years vs -0.5 ($-0.9, -0.2$) at age ≥ 80 years. BP differences were greatest in white people, and especially in women. Triacylglycerol level differences were greatest in younger men. Finally, HbA_{1c} levels were also higher with younger onset diabetes, particularly in black people.

Conclusions/interpretation At diagnosis of type 2 diabetes, when compared with people without diabetes, weight and BP differentials were greater in younger vs older people, in women vs men and in white vs South Asian and black people. These differences were observed even though South Asian and black people tend to develop diabetes a decade earlier with either similar or greater dysglycaemia. These striking patterns may have implications for management and prevention.

Keywords Age · Blood pressure · Ethnicity · HbA_{1c} · Sex · Weight

Darren M. Ashcroft, Martin K. Rutter and Naveed Sattar are joint senior authors.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00125-020-05169-6>) contains peer-reviewed but unedited supplementary material, which is available to authorised users.

✉ Naveed Sattar
naveed.sattar@glasgow.ac.uk

¹ Division of Diabetes, Endocrinology and Gastroenterology, School of Medical Sciences, University of Manchester, Manchester, UK

² Centre for Pharmacoepidemiology and Drug Safety, Division of Pharmacy and Optometry, School of Health Sciences, Manchester Academic Health Sciences Centre, University of Manchester, Manchester, UK

³ Institute of Cardiovascular & Medical Sciences, BHF Glasgow Cardiovascular Research Centre, University of Glasgow, 126 University Place, Glasgow G12 8TA, UK

⁴ Division of Population Health, Health Services & Primary Care, School of Health Sciences, University of Manchester, Manchester, UK

⁵ Department of Biostatistics & Health Informatics, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

⁶ Department of Public Health and Policy, Institute of Population Health Sciences, University of Liverpool, Liverpool, UK

⁷ Manchester Diabetes Centre, Manchester Academic Health Sciences Centre, Manchester University NHS Foundation Trust, Manchester, UK

Research in context

What is already known about this subject?

- Younger people have higher BMIs when they develop type 2 diabetes compared with those developing diabetes at older ages
- Such differences appear more marked in women than men
- South Asian and black people tend to develop diabetes at lower BMIs

What is the key question?

- Are there sex and ethnic differences in risk factors, including BP, lipids and weight, by age of diabetes diagnosis?

What are the new findings?

- The difference in weight between individuals with vs without type 2 diabetes was nearly 20 kg for those diagnosed at age 20–39 years but only around 5 kg for those diagnosed after age 80 years
- Despite lower absolute values, BP excess in diabetes was greatest at a younger age of onset and greatest in white people, especially women
- Triacylglycerol excesses were greater in younger men and HbA_{1c} differences were greater with younger diabetes onset, especially in black people

How might this impact on clinical practice in the foreseeable future?

- Our findings raise awareness of the following: (1) the differing metabolic profiles of people developing type 2 diabetes at different ages, helping explain why younger diabetes onset is more damaging; and (2) a need for greater emphasis on diabetes and CVD management in young people developing diabetes, regardless of their sex or ethnicity

Abbreviations

| | |
|------|-------------------------------------|
| CPRD | Clinical Practice Research Datalink |
| HES | Hospital Episode Statistics |
| IMD | Index of Multiple Deprivation |
| ONS | Office for National Statistics |

Introduction

Recent studies show that diagnosis of type 2 diabetes at younger ages is associated with greater excess CVD risk than when the diagnosis occurs later in life [1, 2]. Diagnosis of type 2 diabetes before the age of 40 years is associated with a 10 year reduction in life expectancy, on average, compared with age-matched people without diabetes. However, this excess risk gradually attenuates when diabetes is diagnosed at older ages, such that there is a negligible effect on life expectancy when diagnosis occurs after 80 years of age [3–5]. There are also historical data suggesting that excess risk of cardiovascular events and all-cause mortality from type 2 diabetes is greater in women than in men [6–9]. We also recently observed that for a given age of diagnosis the number of life-years lost due to type 2 diabetes was lower in South Asian and black people compared with white Europeans [5]. This is somewhat surprising and may reflect earlier CVD risk factor

management in people from minority ethnic groups as they tend to develop diabetes earlier than white people, or it may reflect other poorly understood factors.

There is a paucity of data on characteristics of individuals at the time of diagnosis of diabetes and how this might have changed over time. Some data indicate that clinical characteristics at the time of developing type 2 diabetes vary by age, sex and ethnicity [10–17]. One such variable is BMI, with studies reporting that it is much higher at diabetes diagnosis in younger vs older people, in women vs men, and in people of white European descent vs other ethnicities [15–17]. There are also some data suggesting more adverse lipid profiles and higher BP levels in people who develop diabetes at younger ages vs non-diabetic age- and sex-matched adults, and in women vs men [10–13]. The majority of these studies were limited in their generalisability due to the data source or specific populations considered and they did not include ethnically diverse populations or only observed risk factor levels in people with diabetes.

We hypothesise that at the time of type 2 diabetes diagnosis, CVD risk factors would be higher when compared with those in people without diabetes and that these differences would vary meaningfully by age, sex and ethnicity. We aimed to determine whether differences in risk factors could better explain contrasting risks for developing type 2 diabetes and its

complications. These data are important because they may help focus where more aggressive care is needed to prevent and manage type 2 diabetes.

Methods

Data source We used data from the Clinical Practice Research Datalink (CPRD), an anonymised, longitudinal primary care medical record database from participating UK general practices [18]. Individuals included in the CPRD are broadly representative of the general population in terms of age, sex and ethnicity [18]. The CPRD dataset was linked, at patient-level, to hospitalisation data (Hospital Episode Statistics [HES]), national mortality data (Office for National Statistics [ONS]) and deprivation data (Index of Multiple Deprivation [IMD] 2010) for all eligible patients in 383 linkage-consenting English general practices.

Study population We identified a cohort of individuals with incident diabetes ($N=193,952$), from Read codes (<https://digital.nhs.uk/article/1104/Read-Codes>) in the electronic health record; the cohort included those for whom the first diagnostic code for diabetes (type 1 or type 2 diabetes) was recorded between 1 January 1998 and 31 March 2015, with no diagnoses prior to this date. The index date was taken as the date of the first diabetes diagnostic Read code.

The cohort of individuals identified with incident diabetes were initially assigned to ‘definite’, ‘probable’ and ‘possible’ type 1 diabetes or type 2 diabetes groups. Due to the nature of primary care data and recording, misclassification, misdiagnosis and miscoding can occur [19]. To try and mitigate these errors, we used a validated algorithm for use in routinely collected data that re-classifies type 1 and type 2 diabetes based on diagnostic codes, glucose-lowering drugs, age, BMI and ethnicity [20]. The implementation of the algorithm (including detailed patient numbers for each step) in this population has been described previously [5]. Only those with a final classification of type 2 diabetes ($N=187,968$) were considered (type 1 diabetes $N=5984$). Additionally, people aged less than 20 years at the index date were excluded, resulting in a final type 2 diabetes cohort of 187,601.

Individuals with type 2 diabetes were matched with up to five control individuals without diabetes who were currently registered and contributing data at the cases’ index date, by year of birth (± 2 years), sex, general practice and index date. All individuals were observed from the index date to study end (31 March 2015), practice’s last data collection date, death or transfer out of practice, whichever occurred earliest.

Cohort demographics and baseline characteristics Age, as defined at the index date, was categorised into six groups: 20–39 years; 40–49 years; 50–59 years; 60–69 years; 70–

79 years; and ≥ 80 years. Ethnicity was identified from primary care records and through linkage with HES; the classification of ethnicity using the two data sources has been described previously and further details are provided in electronic supplementary materials (ESM) [Methods](#) and ESM Fig. 1 [5]. Ethnicity was categorised into five groups: white, South Asian, black/black British, other and unknown. Deprivation data was defined using the IMD 2010, a national scheme based on seven deprivation domains and available at small-area level to link with the address of the patient, categorised into five quintiles: IMD 1 (least deprived) to IMD 5 (most deprived) [21]. Drug prescriptions (issue of prescription) at baseline were defined as a prescription within 90 days before or after the index date. Cardiovascular disease (myocardial infarction, stroke, CHD, peripheral vascular disease, angina pectoris) and renal disease (chronic kidney disease stage 4 and above) were defined by Read code, up to the index date.

Outcome measures Body weight, BMI, systolic BP and lipid levels (total cholesterol, HDL-cholesterol, non-HDL-cholesterol, triacylglycerol) were defined as the closest measure up to 12 months before and after the index date (this time window was used to improve capture of available data, particularly for people without diabetes). HbA_{1c} was only examined in people with type 2 diabetes due to the proportion of missing data in those without diabetes.

Statistical analysis Descriptive characteristics for people with type 2 diabetes and matched individuals without diabetes were summarised using mean (\pm SD) and proportions as appropriate.

A multivariate imputation by chained equations (MICE) algorithm was used to impute missing data on baseline variables (BMI, weight, BP, cholesterol, triacylglycerol). Details on the proportion of missingness and differences in characteristics between those with and without missing data are provided in the ESM Tables 1–4 and ESM Key differences in missingness. Imputation models were estimated separately for people with type 2 diabetes and control-group individuals without diabetes within age, age–sex and age–ethnicity strata. See ESM Imputation model for a list of variables included in the imputation model. Five imputed datasets were generated. Analyses comparing biological outcome measures between the type 2 diabetes and control groups were stratified by age, age–sex and age–ethnicity. Multiple linear regression, adjusting for deprivation and accounting for matching, was used to calculate the adjusted mean differences (95% CI) in biological variables between those with and without type 2 diabetes across age groups. Estimates were combined across the five datasets using Rubin’s rules [22]. $p < 0.05$ was considered statistically significant. All statistical analyses were performed using Stata 15.1 (StataCorp, USA).

Table 1 Baseline characteristics of people with incident type 2 diabetes

| Characteristic | All | Sex | | Ethnicity | | |
|--|----------------|---------------|---------------|---------------|-------------|-------------|
| | | Male | Female | White | South Asian | Black |
| <i>N</i> | 187,601 | 103,607 | 83,994 | 143,481 | 9486 | 4451 |
| Men, <i>n</i> (%) | 103,607 (55.2) | | | 78,573 (54.8) | 5142 (54.2) | 2204 (49.5) |
| Age, years | 61.9 ± 14.1 | 60.4 ± 13.4 | 63.7 ± 14.8 | 63.3 ± 13.8 | 52.6 ± 13.5 | 54.0 ± 13.8 |
| Age group, <i>n</i> (%) | | | | | | |
| 20–39 years | 11,997 (6.4) | 6779 (6.5) | 5218 (6.2) | 7170 (5.0) | 1737 (18.3) | 650 (14.6) |
| 40–49 years | 25,770 (13.7) | 15,949 (15.4) | 9821 (11.7) | 16,956 (11.8) | 2310 (24.4) | 1179 (26.5) |
| 50–59 years | 41,491 (22.1) | 25,186 (24.3) | 16,305 (19.4) | 30,162 (21.0) | 2464 (26.0) | 1075 (24.2) |
| 60–69 years | 48,469 (25.8) | 27,917 (27.0) | 20,552 (24.5) | 38,539 (26.9) | 1837 (19.4) | 868 (19.5) |
| 70–79 years | 39,452 (21.0) | 19,876 (19.2) | 19,576 (23.3) | 33,198 (23.1) | 934 (9.9) | 528 (11.9) |
| ≥ 80 years | 20,422 (10.9) | 7900 (7.6) | 12,552 (14.9) | 17,456 (12.2) | 204 (2.2) | 151 (3.4) |
| Ethnicity, <i>n</i> (%) | | | | | | |
| White | 143,481 (76.5) | 78,573 (75.8) | 64,908 (77.3) | | | |
| South Asian | 9486 (5.1) | 5142 (5.0) | 4344 (5.2) | | | |
| Black | 4451 (2.4) | 2204 (2.1) | 2247 (2.7) | | | |
| Other | 2480 (1.3) | 1366 (1.3) | 1114 (1.3) | | | |
| Unknown | 27,703 (14.8) | 16,322 (15.8) | 11,381 (13.6) | | | |
| Deprivation (IMD 2010), <i>n</i> (%) | | | | | | |
| IMD 1 (least) | 34,290 (18.3) | 19,750 (19.1) | 14,540 (17.3) | 26,192 (18.3) | 1238 (13.1) | 222 (5.0) |
| IMD 2 | 40,934 (21.8) | 23,177 (22.4) | 17,757 (21.1) | 32,691 (22.8) | 1405 (14.8) | 365 (8.2) |
| IMD 3 | 38,5337 (20.5) | 21,443 (20.7) | 17,094 (20.4) | 29,504 (20.6) | 1963 (20.7) | 713 (16.0) |
| IMD 4 | 39,128 (20.9) | 21,121 (20.4) | 18,007 (21.4) | 29,432 (20.5) | 2300 (24.3) | 1440 (32.4) |
| IMD 5 (most) | 34,379 (18.3) | 17,956 (17.3) | 16,423 (19.6) | 25,420 (17.7) | 2566 (27.1) | 1703 (38.3) |
| Unknown | 333 (0.2) | 160 (0.2) | 173 (0.2) | 242 (0.2) | 14 (0.2) | 8 (0.2) |
| CVD, <i>n</i> (%) ^a | 37,765 (20.1) | 23,453 (22.6) | 14,312 (17.0) | 32,255 (22.5) | 1191 (12.6) | 359 (8.1) |
| Renal disease (CKD stage ≥4), <i>n</i> (%) | 2335 (1.2) | 1257 (1.2) | 1078 (1.3) | 1943 (1.4) | 109 (1.2) | 55 (1.2) |
| Biological variables (±12 months from diagnosis) | | | | | | |
| BMI, kg/m ² | 31.3 ± 6.9 | 31.1 ± 6.2 | 31.6 ± 7.6 | 31.6 ± 6.9 | 28.5 ± 5.8 | 31.1 ± 6.6 |
| Missing, % | 18.6 | 17.7 | 19.6 | 17.9 | 23.3 | 24.0 |
| Weight, kg | 88.9 ± 21.1 | 94.0 ± 19.9 | 82.4 ± 20.6 | 89.9 ± 21.2 | 77.4 ± 16.3 | 87.7 ± 19.0 |
| Missing, % | 17.0 | 16.2 | 17.9 | 16.4 | 21.4 | 22.4 |
| Total cholesterol, mmol/l | 5.2 ± 1.3 | 5.0 ± 1.3 | 5.3 ± 1.3 | 5.2 ± 1.3 | 5.1 ± 1.2 | 5.1 ± 1.2 |
| Missing, % | 16.8 | 15.9 | 18.0 | 16.2 | 20.6 | 22.4 |
| HDL-cholesterol, mmol/l | 1.2 ± 0.4 | 1.1 ± 0.3 | 1.3 ± 0.4 | 1.2 ± 0.4 | 1.2 ± 0.3 | 1.3 ± 0.4 |
| Missing, % | 28.9 | 28.1 | 29.9 | 28.6 | 28.8 | 30.6 |
| Non-HDL-cholesterol, mmol/l | 4.0 ± 1.3 | 3.9 ± 1.3 | 4.0 ± 1.3 | 3.9 ± 1.3 | 3.9 ± 1.2 | 3.8 ± 1.2 |
| Missing, % | 29.0 | 28.2 | 29.9 | 28.6 | 28.9 | 30.6 |
| Triacylglycerol, mmol/l | 2.6 ± 2.6 | 2.5 ± 3.0 | 2.1 ± 1.9 | 2.3 ± 2.5 | 2.3 ± 3.4 | 1.5 ± 1.5 |
| Missing, % | 29.9 | 28.8 | 31.3 | 29.9 | 28.4 | 29.3 |
| BP | | | | | | |
| Systolic, mmHg | 140 ± 19 | 139 ± 18 | 140 ± 20 | 140 ± 19 | 133 ± 18 | 138 ± 19 |
| Diastolic, mmHg | 81 ± 11 | 81 ± 11 | 80 ± 11 | 81 ± 11 | 81 ± 11 | 82 ± 11 |
| Missing, % | 12.5 | 12.4 | 12.7 | 11.9 | 18.1 | 18.3 |
| HbA _{1c} | | | | | | |
| mmol/mol | 62 ± 22 | 63 ± 23 | 61 ± 21 | 61 ± 22 | 65 ± 22 | 68 ± 27 |
| % | 7.8 ± 2.0 | 7.9 ± 2.1 | 7.7 ± 2.0 | 7.8 ± 2.0 | 8.1 ± 2.0 | 8.4 ± 2.4 |
| Missing (%) | 16.7 | 16.5 | 17.0 | 16.1 | 21.4 | 22.9 |

Table 1 (continued)

| Characteristic | All | Sex | | Ethnicity | | |
|--|------|------|--------|-----------|-------------|-------|
| | | Male | Female | White | South Asian | Black |
| Drug prescriptions, % | | | | | | |
| Diabetes medication | | | | | | |
| Any oral hypoglycaemic | 44.1 | 44.7 | 43.4 | 43.3 | 47.6 | 51.6 |
| Metformin | 34.6 | 35.1 | 33.9 | 33.7 | 39.2 | 40.9 |
| Sulfonylurea | 11.0 | 11.3 | 10.7 | 10.7 | 11.9 | 13.4 |
| Insulin | 3.6 | 3.6 | 3.6 | 3.6 | 3.2 | 6.2 |
| Other | 1.6 | 1.6 | 1.7 | 1.6 | 2.2 | 2.0 |
| Antihypertensive agent | | | | | | |
| Any | 54.9 | 53.0 | 57.2 | 57.8 | 36.5 | 42.8 |
| α-Blocker | 4.3 | 4.5 | 4.0 | 4.6 | 2.3 | 5.1 |
| Angiotensin-2 receptor blocker | 8.7 | 7.7 | 9.9 | 9.1 | 8.0 | 7.5 |
| ACE inhibitor | 29.4 | 31.2 | 27.2 | 30.9 | 19.5 | 20.5 |
| β-Blocker | 19.4 | 19.5 | 19.2 | 21.0 | 11.1 | 10.2 |
| Calcium-channel blocker | 20.6 | 20.5 | 20.7 | 21.2 | 14.4 | 26.5 |
| Diuretic (thiazide, potassium-sparing or loop) | 27.5 | 22.2 | 34.0 | 29.8 | 11.8 | 19.6 |
| Lipid-lowering therapy (any) | 44.1 | 46.0 | 41.8 | 46.0 | 37.1 | 30.6 |
| Antiplatelets (any) | 26.0 | 27.6 | 23.9 | 28.1 | 18.5 | 15.2 |

Data are mean ± SD, *n* (%) or %, where indicated

^a CVD includes myocardial infarction, stroke, CHD, peripheral vascular disease and angina pectoris

CKD, chronic kidney disease

Results

Baseline characteristics of people with type 2 diabetes are presented in Table 1 and those of control groups without diabetes are presented in ESM Table 5. The cohort comprised 187,601 people with incident type 2 diabetes (mean ± SD age 61.9 ± 14.1 years; 55.2% male sex; 76.5% white) and 906,182 individuals without diabetes as matched controls. People with type 2 diabetes were more likely than control individuals to be obese, to have high BP and to be prescribed antihypertensive, lipid-lowering and antiplatelet agents. At diagnosis of type 2 diabetes, women were on average 3 years older than men (Fig. 1a) and had higher mean BMI and cholesterol and lower HbA_{1c}. Women were less likely than men to receive lipid-lowering and antiplatelet medication but were more likely to receive antihypertensive agents (Table 1).

Compared with white people, South Asian and black people were on average ~9–10 years younger at the onset of type 2 diabetes and had a higher level of deprivation (Table 1, Fig. 1b). Black people and particularly South Asian people developed type 2 diabetes at a substantially lower mean weight/BMI than white people and had lower systolic BP levels but higher HbA_{1c}, on average. Correspondingly, compared with white people, a higher percentage of South Asian and black people received prescriptions for glucose-lowering medications and lower proportions were

prescribed lipid-lowering, antihypertensive and antiplatelet medications.

Biological variables at age of type 2 diabetes onset were compared against those for control individuals at the same age without diabetes (ESM Table 6). Younger age at type 2 diabetes diagnosis was associated with higher mean body weight in individuals with diabetes and a greater differential compared with individuals without diabetes. As an illustration, in people aged 20–39 years, the mean ± SD weight was 99.7 ± 27.3 kg for those with type 2 diabetes and 81.0 ± 22.5 kg for those without diabetes (adjusted mean [95% CI] difference 18.7 [18.3, 19.1] kg). However, in people aged ≥80 years, the mean ± SD weight was 72.1 ± 17.0 kg in those with type 2 diabetes and 66.9 ± 18.5 kg for those without diabetes (adjusted mean [95% CI] difference 5.3 [5.0, 5.5] kg). A similar age-related pattern was observed for BMI, systolic BP and triacylglycerol. With the exception of systolic BP levels, higher mean levels of risk factors were observed for people in whom type 2 diabetes was diagnosed at a younger vs older age. Larger differences between those with and without type 2 diabetes were consistently observed in younger age groups, with a narrowing in risk factor levels in older age groups. Whilst this was the case for systolic blood pressure, with the largest difference at ages 20–39 (adjusted mean difference 6 [6, 7] mmHg), in those aged ≥80 years, blood pressure values in those with type 2 diabetes exceeded those in individuals

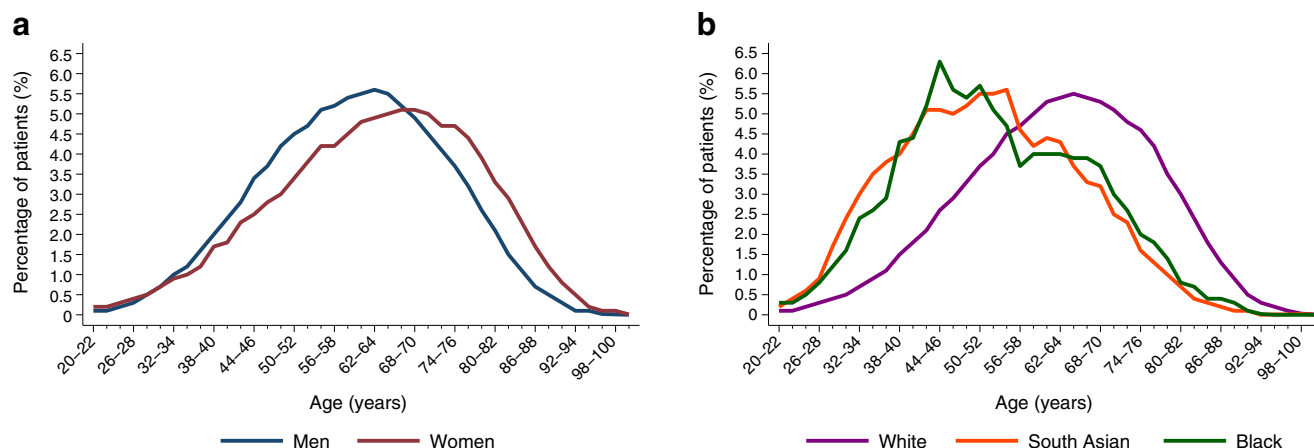


Fig. 1 Age distribution at diagnosis of type 2 diabetes by sex (a) and ethnicity (b)

without diabetes (adjusted mean difference -0.5 [-0.9 , -0.2] mmHg).

This relationship between age and biological variables was observed across sexes. BMI, weight, systolic BP and triacylglycerol differences between those with and without type 2 diabetes were substantially greater in women than in men, particularly at younger ages (Fig. 2a–d and ESM Table 7). For weight, the adjusted mean difference (95% CI) in women aged 20–39 years was 23.1 (22.5, 23.7) kg, reducing to 5.3 (5.0, 5.6) kg in women aged ≥ 80 years. In men, the adjusted mean difference at age 20–39 years was 15.2 (14.6, 15.7) kg and 4.9 (4.5, 5.3) kg at age ≥ 80 years.

Across age groups, there were marked ethnic differences in weight, BMI, BP and lipid levels (ESM Table 8). White people with type 2 diabetes were more likely to have higher levels for these risk factors at diagnosis than South Asian and black people. The differences between those with and without type 2 diabetes was greatest in white people, except for systolic blood pressure where this was lost in those aged ≥ 50 years (Fig. 3a–d and ESM Table 8). With the exception of people aged ≥ 80 years, across all age bands, the mean weight difference (95% CI) between people age 20–39 years with type 2 diabetes and control individuals of white ethnicity was approximately double that observed in South Asian and black people: 24.3 (23.8, 24.9) kg in white people; 11.1 (9.8, 12.3) kg in South Asian people; and 14.2 (11.9, 16.4) kg in black people. Similarly, for BMI at age 20–39 years, the mean difference (95% CI) was 8.4 (8.2, 8.5) kg/m² in white people, 4.1 (3.7, 4.5) kg/m² in South Asian people and 4.9 (4.1, 5.8) kg/m² in black people. For systolic BP and triacylglycerol, the differential between those with and without diabetes was greater in white people than in the other two ethnic groups only up to the ages of 49 years and 69 years, respectively.

While we only examined HbA_{1c} in those with new-onset diabetes, these data showed markedly higher mean HbA_{1c} at younger ages, more so in men than in women and in black

people compared with the two other ethnicities across nearly all age groups (Figs. 2e, 3e).

The same patterns were observed across ethnic–sex subgroups (ESM Table 9). The impact of diabetes on BMI, weight and systolic BP was most substantial in white women at any given age compared with South Asian and black men and women and was particularly prominent in younger white women. At age 20–39 years, the mean (95% CI) difference in weight between individuals with and without type 2 diabetes was 27.6 (26.9, 28.3) kg in white women, 13.1 (11.4, 14.7) kg in South Asian women, 16.6 (14.2, 19.0) kg in black women, 20.2 (19.5, 20.9) kg in white men, 7.6 (5.7, 9.4) kg in South Asian men and 11.1 (7.6, 14.7) kg in black men. Similarly, at age 20–39 years, the mean (95% CI) difference in systolic BP between those with and without type 2 diabetes was 9.7 (9.0, 10.3) mmHg in white women, 7.6 (5.7, 9.5) mmHg in South Asian women, 6.8 (4.0, 9.6) mmHg in black women, 7.6 (6.9, 8.3) mmHg in white men, 5.2 (3.7, 6.7) mmHg in South Asian men and 6.1 (3.6, 8.6) mmHg in black men.

Discussion

In this large contemporary primary care-based study from England, our findings were as follows: (1) there was an association between younger age at type 2 diabetes diagnosis and a higher number of elevated cardiovascular risk factors; (2) the differences in weight, BMI and systolic BP in individuals with vs without type 2 diabetes were more marked in women than in men, especially at younger ages, whereas HbA_{1c} and triacylglycerol tended to be higher in men; (3) white people with type 2 diabetes were more likely to have higher weight, BMI, BP and triacylglycerol levels across all age groups (except for the very oldest groups) than South Asian and black people with type 2 diabetes. This poorer cardiovascular risk factor profile may help explain previous observations that young

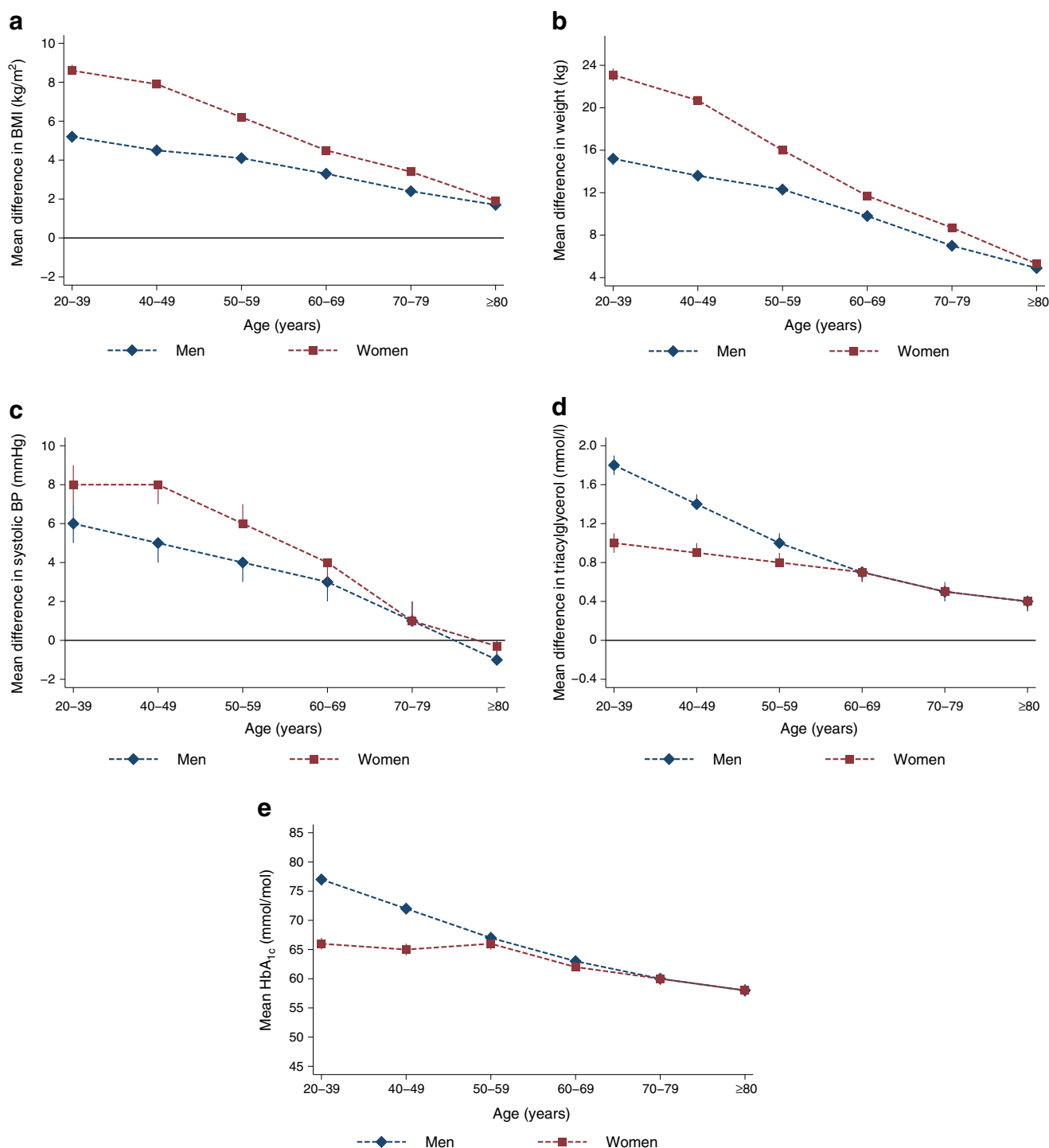


Fig. 2 (a–d) Adjusted age-specific mean (95% CI) differences in BMI (a), weight (b), systolic BP (c) and triacylglycerol level (d) in men and women recently diagnosed with type 2 diabetes compared with men and

women without diabetes. (e) Age-specific mean HbA_{1c} levels in men and women recently diagnosed with type 2 diabetes

people diagnosed with type 2 diabetes, across all ethnicities, lose more life-years from diabetes than those diagnosed when older [5]. It also explains why, in historical studies, women had greater increase in their relative risk for CVD than men when they develop diabetes, with high BP being a stronger risk factor for CVD than raised triacylglycerol [5, 23]. These

data also might explain our prior finding which linked greater loss of life-years from type 2 diabetes in white vs South Asian and black people [5].

Regarding sex differences, it has been established that women generally have a higher BMI than men when diagnosed with type 2 diabetes [14, 24, 25]. In both men and

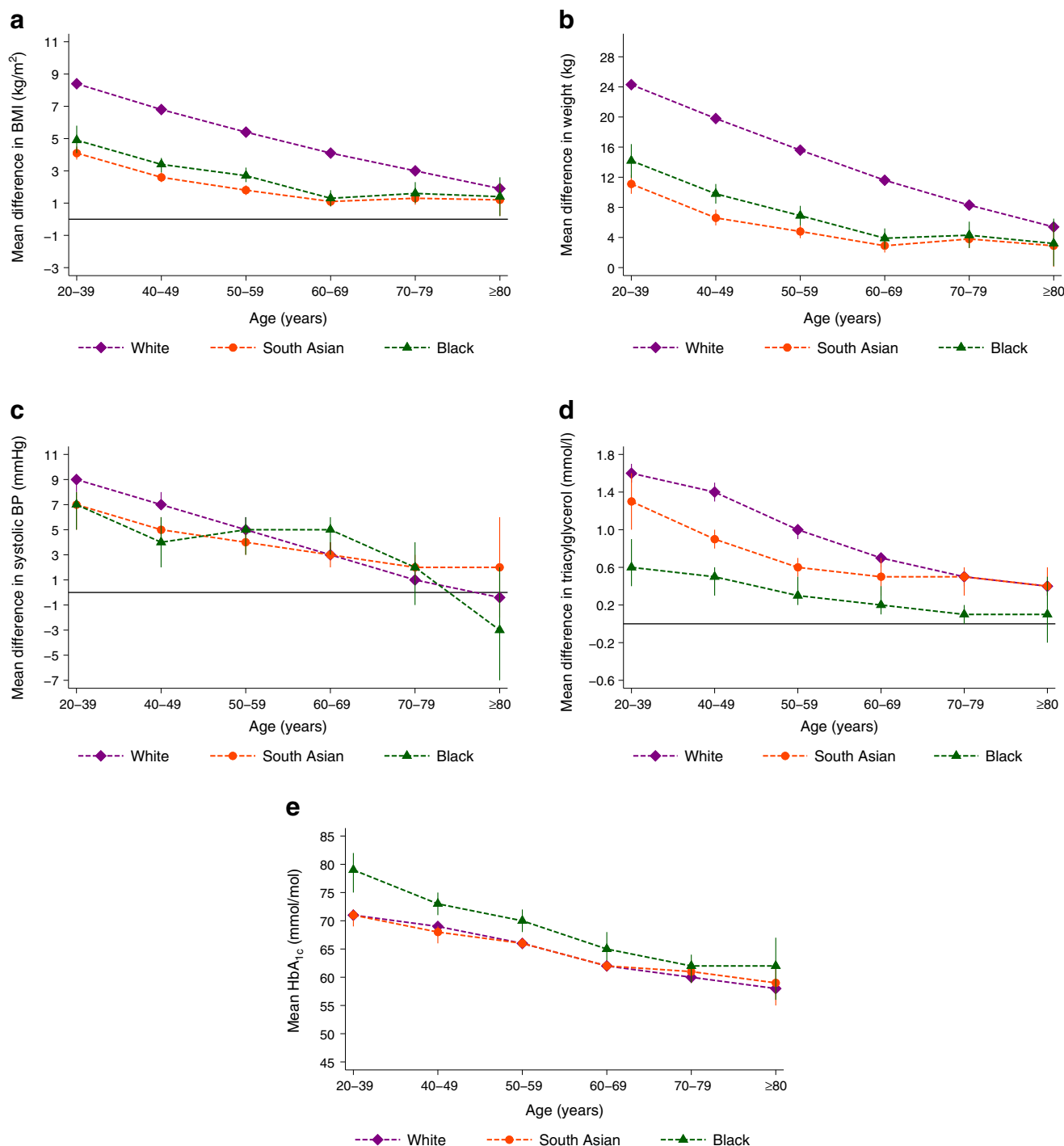


Fig. 3 Adjusted age-specific mean (95% CI) differences in BMI (a), weight (b), systolic BP (c) and triacylglycerol level (d) in white, South Asian and black people recently diagnosed with type 2 diabetes compared

with people without diabetes. (e) Age-specific mean HbA_{1c} levels in white, South Asian and black people recently diagnosed with type 2 diabetes $\geq m^2$

women, the BMI differences are most marked at younger ages of diagnosis and the differences narrow with diagnosis at older ages. A younger age of onset of obesity leads to a greater cumulative exposure to obesity, which may contribute to a younger age of onset of diabetes [15, 26]. Older people being diagnosed with diabetes at a lower body weight could in part

be explained by them having a lower cumulative lifetime exposure to obesity.

Young women appear to undergo larger excess weight gain than men prior to being diagnosed with type 2 diabetes [24, 25]. It has been suggested that women may experience greater adverse metabolic changes than men as they develop diabetes,

with a poorer cardiovascular risk factor profile when compared with their non-diabetic counterparts. However, while this has been well documented for BMI, weight differences have not been reported and less is known about the impact on other risk factors, especially across the age spectrum. In our comprehensive comparison of seven cardiovascular risk markers, we identified bigger differences (compared with their non-diabetic counterparts) in weight, systolic BP and HDL-cholesterol in women than in men at the diagnosis of type 2 diabetes at any age but particularly in younger people. The converse was seen with triacylglycerol, with men showing a trend for higher levels at the onset of type 2 diabetes, most notably in the youngest age groups. Understanding these patterns, particularly the sex differences, are important considerations for future cardiovascular risk evaluation. Finally, sex biases in prescribing of medications to manage cardiovascular risk factors in people with diabetes have been identified: women are less likely to be prescribed medications for adverse risk factors and/or established CVD [27–30]. Therefore, greater awareness of sex- and age-related differences in biological factors should prompt appropriate intervention.

Regarding ethnic differences, the South Asian and black people with type 2 diabetes were on average 9–10 years younger than the white population. This may arise from the higher diabetes incidence rates seen in younger age groups or may reflect the age distribution of the different ethnic groups. South Asian and black people were observed to present with a different cardiometabolic risk profile, with typically higher HbA_{1c}, lower BMI, and established micro- and macrovascular complications [14, 31–33]. As identified previously by Paul et al. [14], and shown in this cohort, white people with type 2 diabetes had a pattern of presenting with higher BMI at any age compared with South Asian and black people, with larger ethnic differences in BMI levels reported at younger ages. We extend these findings by observing a similar pattern across other risk factors including weight, systolic BP and triacylglycerol in white people, as well as a tendency towards higher HbA_{1c} levels in black people. Furthermore, greater risk factor burden at diagnosis was seen in white people, suggesting that they may undergo a greater deterioration in cardiometabolic risk factors on their pathway to developing diabetes. In particular, white women with type 2 diabetes had the most substantial burden of CVD risk factors when compared with their non-diabetic white counterparts. These findings may, in part, help to explain mechanisms behind findings from our previous research on this cohort wherein white people and particularly white women with type 2 diabetes had a greater likelihood for poorer life expectancy and a greater risk of CVD than South Asian and black women [5]. Earlier treatment of abnormal cardiometabolic risk factors in South Asian and black people may also play a role, given their development of disease at a younger age. As South Asian people are known to have an elevated risk of CVD

compared with white Europeans [34], there has been a greater emphasis in these populations, particularly for those with type 2 diabetes, to reduce modifiable risk factors through lifestyle changes and drug management for primary and secondary CVD prevention [35].

This study has several important strengths: (1) we obtained a large cohort of people with type 2 diabetes identified from primary care, reducing the likelihood of significant selection bias; (2) we applied a validated algorithm to identify type 2 diabetes to mitigate any potential miscoding or misclassification from electronic health records; (3) we had clinical and prescribing information on people with and without type 2 diabetes; and (4) we combined data from general practice and hospital records to increase completeness and accuracy of ethnicity information.

Our study has several potential limitations. First, there is the potential for misclassification of diabetes when using routinely collected primary care data without validation from consultation free text. The accuracy of primary care Read codes depends on the team entering them, the clinicians' time and information-technology skills, certainty of diagnosis and organisational issues [36]. The algorithm proposed and validated by de Lusignan et al. [20] uses additional information beyond Read codes alone. These clinical features have been identified by The Royal College of General Practitioners and others, as being important for differentiating between type 1 and type 2 diabetes [19, 37]; therefore, after applying the algorithm to our cohort, we would expect the prevalence of misclassification of the final diabetes type to be lower than the initial classification. However, any residual misclassification might bias our results towards the null, so our results are likely to be conservative estimates of the true difference between control individuals and people diagnosed with type 2 diabetes. Second, depending on the characteristics of an individual, the assessment of type 2 diabetes prompted by a clinician may represent either screening or diagnostic testing [38, 39]. These are two distinct pathways to determining diabetes status and may result in observations of individuals at different stages of the condition at the time of diagnosis. Third, prescription data in the CPRD only indicate when a prescription was issued but not whether the medication was dispensed or taken as recommended. Fourth, there was a large proportion of missing data for cardiovascular risk factors in people without diabetes. We addressed this through multiple imputation with the imputation model including an extensive range of patient characteristics and auxiliary variables; importantly, we noted no clinically meaningful differences in measurements of biological variables when comparing those with and without any missing data (ESM Tables 2–4). We acknowledge that missing data may have occurred preferentially in people with normal or close to normal risk factors (i.e. missing not at random), possibly leading to bias in the magnitude of the association and an underestimation of the differentials in risk factor levels. Fifth,

people without type 2 diabetes and attending primary care for cardiovascular risk factor assessment are likely to have abnormal cardiovascular risk factors or be at risk of diabetes or CVD. Therefore, measured differences between those with and without diabetes are potentially underestimated when compared with a healthier non-diabetic control group. Sixth, due to the availability of clinical data, we allowed for the baseline risk factors to be defined from data captured up to 12 months after the index date. This reduced the proportion of missing data in people without diabetes who may not have measures frequently recorded. The majority of people diagnosed with diabetes will have had information recorded either on or close to this date. However, a small proportion of individuals with diabetes would have risk factors recorded only after clinical intervention. Such intervention would be expected to improve but not normalise abnormal risk factors. Therefore, the observed risk factor differences compared with control counterparts may be somewhat underestimated by our analysis. Seventh, linkage with hospital data is only available for people attending English general practices; therefore, generalisation of our findings to other healthcare systems, other countries and other ethnic groups at high risk of diabetes may be limited. However, we would anticipate broadly similar findings. Finally, we recognise that less frequent screening means diabetes is being diagnosed later in the course of disease in younger people. Notably, this cannot explain BMI differentials by sex (men had higher HbA_{1c} levels but lower weight differences than women) or by ethnicity (HbA_{1c} levels were higher or the same but weight differences were far smaller). It is also unlikely to explain the substantial weight differentials seen across the life course.

Targeted screening and active modification of weight in men and South Asian and black people, who may be at risk of developing type 2 diabetes at lower BMI values, may be warranted to prevent the development of type 2 diabetes. Active intervention to reduce cardiovascular risk factors other than weight is justified in all people with type 2 diabetes and those at high risk for diabetes but these data argue for particular focus in younger people, where perhaps more aggressive risk factor management may be needed despite lower mean blood pressures and lower ten year CVD risks, but high lifetime risks. Further clinical trials are required to assess the degree of weight reduction required to achieve diabetes remission by age and ethnicity.

In conclusion, we have provided perhaps the most comprehensive assessment of age-, sex- and ethnicity-based differences in established CVD risk factors at the diagnosis of type 2 diabetes in a high-income country. The findings may help explain why younger people are more likely to lose additional life-years when they develop diabetes and also perhaps the contrast between outcomes in white people and other ethnicities. Clinically, given that men and people from minority ethnic groups tend to develop diabetes at far lower BMI values, better targeting of these populations to prevent overt

type 2 diabetes is warranted. However, in all ethnic groups, risk factor differences, as compared with people without diabetes, are worse at younger ages suggesting a much greater need to improve diabetes care in younger people, regardless of their sex or ethnicity. The high prevalence of abnormal CVD risk factors in younger people developing diabetes suggests that there may be value in performing trials in these individuals evaluating the impact of more aggressive lifestyle and/or pharmacotherapy to promote weight loss and CVD risk reduction.

Data availability Read codes are publicly available at The ClinicalCodes repository [40] and can be accessed at <https://clinicalcodes.rss.mhs.man.ac.uk/>. Electronic health records are, by definition, considered 'sensitive' data in the UK by the Data Protection Act and cannot be shared via public deposition because of information governance restriction in place to protect patient confidentiality. Access to data is available only once approval has been obtained through the individual constituent entities controlling access to the data. The primary care data can be requested via application to the Clinical Practice Research Datalink (<https://www.cprd.com>); secondary care data can be requested via application to the Hospital Episode Statistics from the UK Health and Social Care Information Centre (www.hscic.gov.uk/hesdata); and mortality data are available by application to the UK Office for National Statistics (www.ons.gov.uk/ons/index.html).

Funding This study was funded by Diabetes UK (BDA: 14/0004971). The views expressed are those of the authors and not necessarily those of Diabetes UK. The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review or approval of the manuscript and decision to submit the manuscript for publication.

Authors' relationships and activities NS reports grants and personal fees from Boehringer Ingelheim and personal fees from Amgen, AstraZeneca, Eli Lilly, Janssen, Novo Nordisk and Sanofi, outside the submitted work. DMA reports research funding from AbbVie, Almirall, Celgene, Eli Lilly, Novartis, Union Chimique Belge (UCB) and the LEO Foundation outside the submitted work. MKR reports receiving research funding from Novo Nordisk, consultancy fees from Novo Nordisk and Roche Diabetes Care, and modest owning of shares in GlaxoSmithKline, outside the submitted work. All other authors declare that there are no relationships or activities that might bias, or be perceived to bias, their work.

Contribution statement The idea for this paper was generated by NS, AKW, DMA and MKR. The analyses were done by AKW with support from NS, MKR and DMA. All authors contributed to the interpretation of the data, drafting and critical revision of the manuscript and final approval for publication. NS is guarantor and takes responsibility for the contents of the article.

Ethics approval This study is based in part on data from the CPRD obtained under licence from the UK Medicines and Healthcare products Regulatory Agency. The data are provided by patients and collected by the NHS as part of their care and support. ONS and HES data are subject to Crown copyright (2018) protection, re-used with the permission of The Health, & Social Care Information Centre, all rights reserved. The OPCS Classification of Interventions and Procedures, codes, terms and text is Crown copyright (2016) published by the Health and Social Care Information Centre, also known as NHS Digital, and licensed under the Open Government Licence available at www.nationalarchives.gov.uk/doc/open-government-licence/open-government-licence.htm. The interpretation and conclusions contained in this study are those of the

authors alone. The study was approved by the Independent Scientific Advisory Committee (ISAC) for CPRD research, ref. 15_123Mn.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Al-Saeed AH, Constantino MI, Molyneaux L et al (2016) An inverse relationship between age of type 2 diabetes onset and complication risk and mortality: the impact of youth-onset type 2 diabetes. *Diabetes Care* 39(5):823–829. <https://doi.org/10.2337/dc15-0991>
- Huo L, Magliano DJ, Rancière F et al (2018) Impact of age at diagnosis and duration of type 2 diabetes on mortality in Australia 1997–2011. *Diabetologia* 61(5):1055–1063. <https://doi.org/10.1007/s00125-018-4544-z>
- Morgan CL, Currie CJ, Peters JR (2000) Relationship between diabetes and mortality: a population study using record linkage. *Diabetes Care* 23(8):1103–1107. <https://doi.org/10.2337/diacare.23.8.1103>
- Walker J, Colhoun H, Livingstone S et al (2018) Type 2 diabetes, socioeconomic status and life expectancy in Scotland (2012–2014): a population-based observational study. *Diabetologia* 61(1):108–116. <https://doi.org/10.1007/s00125-017-4478-x>
- Wright AK, Kontopantelis E, Emsley R et al (2017) Life expectancy and cause-specific mortality in type 2 diabetes: a population-based cohort study quantifying relationships in ethnic subgroups. *Diabetes Care* 40(3):338–345. <https://doi.org/10.2337/dc16-1616>
- Peters SAE, Huxley RR, Woodward M (2014) Diabetes as risk factor for incident coronary heart disease in women compared with men: a systematic review and meta-analysis of 64 cohorts including 858,507 individuals and 28,203 coronary events. *Diabetologia* 57(8):1542–1551. <https://doi.org/10.1007/s00125-014-3260-6>
- Peters S, Huxley R, Woodward M (2014) Diabetes as a risk factor for stroke in women compared with men: a systematic review and meta-analysis of 64 cohorts, including 775,385 individuals and 12,539 strokes. *Lancet* 383(9933):1973–1980. [https://doi.org/10.1016/S0140-6736\(14\)60040-4](https://doi.org/10.1016/S0140-6736(14)60040-4)
- Huxley R, Barzi F, Woodward M (2006) Excess risk of fatal coronary heart disease associated with diabetes in men and women: meta-analysis of 37 prospective cohort studies. *BMJ* 332(7533):73–78. <https://doi.org/10.1136/bmj.38678.389583.7C>
- Prospective Studies Collaboration and Asia Pacific Cohort Studies Collaboration (2018) Sex-specific relevance of diabetes to occlusive vascular and other mortality: a collaborative meta-analysis of individual data from 980 793 adults from 68 prospective studies. *Lancet Diabetes Endocrinol* 6(7):538–546
- Nanayakkara N, Ranasinha S, Gadowski AM et al (2018) Age-related differences in glycaemic control, cardiovascular disease risk factors and treatment in patients with type 2 diabetes: a cross-sectional study from the Australian National Diabetes Audit. *BMJ Open* 8(8):e020677. <https://doi.org/10.1136/bmjopen-2017-020677>
- Khang AR (2018) Letter: clinical characteristics of people with newly diagnosed type 2 diabetes between 2015 and 2016: difference by age and body mass index (*Diabetes Metab J* 2018;42:137–46). *Diabetes Metab J* 42(3):249–250. <https://doi.org/10.4093/dmj.2018.0074>
- Steinarsson AO, Rawshani A, Gudbjörnsdóttir S, Franzén S, Svensson A-M, Sattar N (2018) Short-term progression of cardiometabolic risk factors in relation to age at type 2 diabetes diagnosis: a longitudinal observational study of 100,606 individuals from the Swedish National Diabetes Register. *Diabetologia* 61(3):599–606. <https://doi.org/10.1007/s00125-017-4532-8>
- Schroeder EB, Bayliss EA, Daugherty SL, Steiner JF (2014) Gender differences in cardiovascular risk factors in incident diabetes. *Women's Health Issues* 24(1):e61–e68. <https://doi.org/10.1016/j.whi.2013.09.008>
- Paul SK, Owusu Adjah ES, Samanta M et al (2017) Comparison of body mass index at diagnosis of diabetes in a multi-ethnic population: a case-control study with matched non-diabetic controls. *Diabetes Obes Metab* 19(7):1014–1023. <https://doi.org/10.1111/dom.12915>
- Chen Y, Zhang X-P, Yuan J et al (2018) Association of body mass index and age with incident diabetes in Chinese adults: a population-based cohort study. *BMJ Open* 8(9):e021768. <https://doi.org/10.1136/bmjopen-2018-021768>
- Logue J, Walker JJ, Colhoun HM et al (2011) Do men develop type 2 diabetes at lower body mass indices than women? *Diabetologia* 54(12):3003–3006. <https://doi.org/10.1007/s00125-011-2313-3>
- Kwon SK (2014) Women are diagnosed with type 2 diabetes at higher body mass indices and older ages than men: Korea National Health and Nutrition Examination Survey 2007–2010. *Diabetes Metab J* 38(1):74–80. <https://doi.org/10.4093/dmj.2014.38.1.74>
- Herrett E, Gallagher AM, Bhaskaran K et al (2015) Data resource profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol* 44(3):827–836. <https://doi.org/10.1093/ije/dyv098>
- Royal College of General Practitioners (2011) Coding, classification and diagnosis of diabetes. Available from https://clininf.eu/wp-content/uploads/2017/02/nhs_diabetes_and_rcgp_cod_final_report.pdf. Accessed 20 Mar 2019
- de Lusignan S, Khunti K, Belsey J et al (2010) A method of identifying and correcting miscoding, misclassification and misdiagnosis in diabetes: a pilot and validation study of routinely collected data. *Diabet Med* 27(2):203–209. <https://doi.org/10.1111/j.1464-5491.2009.02917.x>
- McLennan D, Barnes H, Noble M, Dibben C (2011) The English indices of deprivation 2010. Available from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/6320/1870718.pdf. Accessed 10 Mar 2016
- Rubin DB (1996) Multiple imputation after 18+ years. *J Am Stat Assoc* 91(434):473–489. <https://doi.org/10.1080/01621459.1996.10476908>
- JBS3 Board (2014) Joint British Societies' consensus recommendations for the prevention of cardiovascular disease (JBS3). *Heart* 100(Suppl 2):ii1–ii67
- Sattar N (2013) Gender aspects in type 2 diabetes mellitus and cardiometabolic risk. *Best Pract Res Clin Endocrinol Metab* 27(4):501–507. <https://doi.org/10.1016/j.beem.2013.05.006>
- Wannamethee SG, Papacosta O, Lawlor DA et al (2012) Do women exhibit greater differences in established and novel risk factors between diabetes and non-diabetes than men? The British Regional Heart Study and British Women's Heart Health Study. *Diabetologia* 55(1):80–87. <https://doi.org/10.1007/s00125-011-2284-4>

26. Wei GS, Coady SA, Reis JP et al (2015) Duration and degree of weight gain and incident diabetes in younger versus middle-aged black and white adults: ARIC, CARDIA, and the Framingham Heart Study. *Diabetes Care* 38(11):2042–2049. <https://doi.org/10.2337/dc14-2770>
27. Wright AK, Kontopantelis E, Emsley R et al (2019) Cardiovascular risk and risk factor management in type 2 diabetes mellitus: a population-based cohort study assessing sex disparities. *Circulation* 139(24):2742–2753. <https://doi.org/10.1161/CIRCULATIONAHA.118.039100>
28. Wexler D, Grant R, Meigs J, Nathan D, Cagliero E (2005) Sex disparities in treatment of cardiac risk factors in patients with type 2 diabetes. *Diabetes Care* 28(3):514–520. <https://doi.org/10.2337/diacare.28.3.514>
29. Vimalananda VG, Miller DR, Palnati M, Christiansen CL, Fincke BG (2011) Gender disparities in lipid-lowering therapy among veterans with diabetes. *Womens Health Issues* 21(4):S176–S181. <https://doi.org/10.1016/j.whi.2011.04.009>
30. Yu MK, Lyles CR, Bent-Shaw LA, Young BA (2013) Sex disparities in diabetes process of care measures and self-care in high-risk patients. *J Diabetes Res* 2013:1–8. <https://doi.org/10.1155/2013/575814>
31. Misra A, Sattar N, Tandon N et al (2018) Clinical management of type 2 diabetes in south Asia. *Lancet Diabetes Endocrinol* 6(12):979–991. [https://doi.org/10.1016/S2213-8587\(18\)30199-2](https://doi.org/10.1016/S2213-8587(18)30199-2)
32. Dal Canto E, Farukh B, Faconti L (2018) Why are there ethnic differences in cardio-metabolic risk factors and cardiovascular diseases? *JRSM Cardiovasc Dis* 7:1–5
33. Johns E, Sattar N (2017) Cardiovascular and mortality risks in migrant South Asians with type 2 diabetes: are we winning the battle? *Curr Diab Rep* 17(10):100. <https://doi.org/10.1007/s11892-017-0929-5>
34. Lip GYH, Barnett AH, Bradbury A et al (2007) Ethnicity and cardiovascular disease prevention in the United Kingdom: a practical approach to management. *J Hum Hypertens* 21(3):183–211. <https://doi.org/10.1038/sj.jhh.1002126>
35. The National Institute for Health and Care Excellence (NICE) (2017) Cardiovascular disease risk assessment and prevention. Available from <https://bnf.nice.org.uk/treatment-summary/cardiovascular-disease-risk-assessment-and-prevention.html>. Accessed 20 Jan 2019
36. Tate AR, Dungey S, Glew S, Beloff N, Williams R, Williams T (2017) Quality of recording of diabetes in the UK: how does the GP's method of coding clinical data affect incidence estimates? Cross-sectional study using the CPRD database. *BMJ Open* 7(1):e012905. <https://doi.org/10.1136/bmjopen-2016-012905>
37. Shields BM, Peters JL, Cooper C et al (2015) Can clinical features be used to differentiate type 1 from type 2 diabetes? A systematic review of the literature. *BMJ Open* 5(11):e009088. <https://doi.org/10.1136/bmjopen-2015-009088>
38. National Institute for Health and Care Excellence (2017) Type 2 diabetes: prevention in people at high risk. Public health guideline [PH38]. Available from <https://www.nice.org.uk/guidance/ph38/chapter/Recommendations>. Accessed 20 Mar 2019
39. American Diabetes Association (2003) Screening for type 2 diabetes. *Diabetes Care* 26(Suppl 1):S21–S24
40. Springate DA, Kontopantelis E, Ashcroft DM et al (2014) ClinicalCodes: an online clinical codes repository to improve the validity and reproducibility of research using electronic medical records. *PLoS One* 9(6):e99825. <https://doi.org/10.1371/journal.pone.0099825>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.